Correspondence

The double-edged nature of using genetic databases: melanotransferrin genes and transcripts

Eric Sekyere, Louise L. Dunn, Des R. Richardson*

First published online 16 June 2003

Early access to the human genome database provides tremendous advantages in accelerating biomedical research. It has made available huge amounts of DNA sequence that aids the research efforts of countless scientists. The drawback of this approach is the inherent risk of transferring errors from the assembled sequences into published work.

Melanotransferrin is a membrane-bound molecule that shares homology to the serum Fe-binding protein, transferrin [1]. However, while this molecule can bind Fe through a specific Fe-binding site, its precise function remains unknown [1]. Based on the genomic sequences in the GenBank database, in 2002 we reported the identification of a second melanotransferrin gene (MTf2) and a new short MTf transcript or splice variant (e.g. NM_033316) [2].

More recent revision of the database indicates that the contig on which the MTf gene was mapped (NT_005834) is now considered 'finished', as this region has significant experimental data for its validation (Rana C. Morris, NCBI User Services, personal communication). Considering this, we again examined the contig, and found that the MTf2 gene no longer existed in the validated version of the human genome sequence. Therefore, the presence of the MTf2 gene in the working draft of the database appeared to be an unavoidable error generated during collating countless genomic sequences. Hence, while it is important to have access to the genome due to its potential for advancing biomedical research [3], our experience highlights the potential pitfalls in this approach.

On the other hand, the existence of the reported short MTf splice variant (e.g. NM_033316) has been confirmed in our laboratory using RT-PCR and sequencing in SK-Mel-28 melanoma cells. Initial work from Northern blotting studies of normal tissues together with further RT-PCR results indicate that this second transcript may be differentially expressed. Obviously, the differential expression of this transcript may be important in terms of understanding the function of MTf. Hence, in this case, access to unrefereed genetic databases was very useful in promoting our research effort. In conclusion, our studies with MTf demonstrate the double-edged nature of early versions of the genetic database and the caveats involved in its use.

References

[1] Sekyere, E. and Richardson, D.R. (2000) FEBS Lett. 483, 11–16.
[2] Sekyere, E., Food, M.R. and Richardson, D.R. (2002) FEBS Lett. 512, 350–352.
[3] Jang, W., Chen, H-C., Sicotte, H. and Schuler, G.D. (1999) Trends Genet. 15, 284–286.

*Corresponding author. Fax: (61)-2-9382 1850.
E-mail address: d.richardson@ccia.org.au (D.R. Richardson).

Children's Cancer Institute Australia for Medical Research, Iron Metabolism and Chelation Program, P.O. Box 81, High St., Randwick, Sydney, NSW 2031, Australia

0014-5793/03/$22.00 © 2003 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(03)00658-6